# A Novel Symmetric Skew-Hamiltonian Isotropic Lanczos Algorithm for Spectral Conformal Parameterizations

Wei-Qiang Huang · Xianfeng David Gu · Wen-Wei Lin · Shing-Tung Yau

Received: 27 October 2013 / Revised: 27 February 2014 / Accepted: 4 March 2014 / Published online: 15 March 2014 © Springer Science+Business Media New York 2014

Abstract In the past decades, many methods for computing conformal mesh parameterizations have been developed in response to demand of numerous applications in the field of geometry processing. Spectral conformal parameterization (SCP) (Mullen et al. in Proceedings of the symposium on geometry processing, SGP '08. Eurographics Association, Aire-la-Ville, Switzerland, pp 1487-1494, 2008) is one of these methods used to compute a quality conformal parameterization based on the spectral techniques. SCP focuses on a generalized eigenvalue problem (GEP)  $L_C \mathbf{f} = \lambda B \mathbf{f}$  whose eigenvector(s) associated with the smallest positive eigenvalue(s) provide the conformal parameterization result. This paper is devoted to studying a novel eigensolver for this GEP. Based on structures of the matrix pair  $(L_C, B)$ , we show that this GEP can be transformed into a small-scale compressed and deflated standard eigenvalue problem with a symmetric positive definite skew-Hamiltonian operator. We then propose a symmetric skew-Hamiltonian isotropic Lanczos algorithm (SHILA) to solve the reduced problem. Numerical experiments show that our compressed deflating technique can exclude the impact of convergence from the kernel of  $L_{C}$  and transform the original problem to a more robust system. The novel SHILA method can effectively avoid the disturbance of duplicate eigenvalues. As a result, based on the spectral model of SCP, our numerical eigensolver can compute the conformal parameterization accurately and efficiently.

W.-W. Lin e-mail: wwlin@math.nctu.edu.tw

X. D. Gu Department of Computer Science, Stony Brook University, Stony Brook, NY 11794-4400, USA e-mail: gu@cs.sunysb.edu

#### S.-T. Yau Department of Mathematics, Harvard University, Cambridge, MA 02138, USA e-mail: yau@math.harvard.edu

W.-Q. Huang (⊠) · W.-W. Lin Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan e-mail: wqhuang@math.nctu.edu.tw

**Keywords** Conformal parameterization · Nonequivalence deflation · Null-space free · Symmetric skew-Hamiltonian · Isotropic Lanczos

Mathematics Subject Classification 15B57 · 65D18 · 65F15 · 68U05

### 1 Introduction

Matrix computation is a fundamental tool in digital geometry processing. Many interesting and challenging problems in computational geometry are eventually confronted with the difficulty of how to solve the corresponding problems within the context of matrix computation, such as linear systems, eigenvalue problems, optimization problems, and so on. Certainly, there have been many excellent theoretical investigations as well as numerical algorithms for these subjects, and numerous well-developed libraries, such as LAPACK [2], ARPACK [24], PETSc [4], SPLEPc [17], etc., are capable of handling these problems. Nevertheless, the existing packages may still encounter some difficulties such as getting erratic or unwanted solutions, suffering from slow convergence, or even failing to converge. Therefore, it is still an important issue to study accurate, efficient and robust numerical algorithms as well as mathematical analyses specially tailored to take advantage of the structures and properties of the target problem. In this paper, we focus on a generalized eigenvalue problem (GEP) arising from the *spectral conformal parameterization* (SCP) [31], which is one of those free-boundary and angle-preserving mesh parameterizations.

Mesh parameterization is an important and active subject in the research of digital geometry processing. Its goal is to construct a piecewise linear map between a triangulated 3D mesh surface and a 2D planar mesh. Once we obtain appropriate parameterizations, any complicated processing tasks on surfaces can be transformed into easier ones on the planar domain through the correspondences of geometric information between the surface mesh and the planar mesh. Mesh parameterizations unavoidably introduce distortion in angles, areas or lengths, and the main challenge for parameterization approaches is to minimize the resulting distortion in some sense as much as possible. Maps that preserve the angle are called *conformal* maps; that preserve the area are called *authalic* maps and that preserve the length are called *isometric* maps. It is known that a mapping, which is both conformal and authalic, must be isometric. Excellent surveys on various kinds of mesh parameterization techniques can be found in [8,20,37], and the references therein. Many feasible conformal parameterization methods have been intensively studied and developed ever since several applications required angle-preserving parameterizations. Such applications include texture mapping, remeshing, compression, recognition, and morphing, to name just a few. According to the outputs of conformal parameterizations, most of them can be classified into one of the following categories: map category [6,9,12,18,25,31], differential 1-form category [11, 14, 22, 34], angle structure category [23, 36, 38, 42], and metric category [21, 41]. For a comprehensive survey on the topic of theoretical and computational conformal geometry, we refer to [13, 15, 16]. Here we only briefly review the most related works to motivate our research in this paper.

The discrete conformal parameterization (DCP) proposed by Desbrun et al. [6] computes the conformal parameterization by minimizing the Dirichlet energy defined on triangular meshes subject to so-called natural boundary conditions. Through the least-squares approximation of the discrete Cauchy–Riemann equation, Lévy et al. [25] introduced the approach of *least squares conformal maps* (LSCM) for computing conformal mesh parameterization. These two methods can achieve parameterizations with much lower angle distortion and, as shown in [5], LSCM and DCP are theoretically equivalent. More recently, Mullen et al. [31] presented a spectral approach, named SCP, to reduce common artifacts of LSCM and DCP due to positional constraints or mesh sampling irregularity, and thereby achieve highquality conformal parameterizations. SCP tends to compute the conformal parameterization via a constrained energy minimization problem which can be transformed into to a GEP  $L_C \mathbf{f} = \lambda B \mathbf{f}$  [31] with  $(L_C, B)$  being a symmetric positive semi-definite matrix pair. Therefore, determining the minimizer of the constrained optimization problem, i.e., the desired result of the parameterization, is equivalent to finding the smallest positive eigenvalue  $\lambda$  and the associated eigenvector  $\mathbf{f}$  of  $(L_C, B)$ .

For the computation of the GEP  $L_C \mathbf{f} = \lambda B \mathbf{f}$ , Mullen et al. [31] and, most recently, Alexa and Wardetzky [1] individually proposed feasible numerical methods. The former considered an inverted modified eigenvalue problem instead of the original one; the latter reformulated the original problem as an equivalent small-scale problem. However, when processing large mesh models, these two approaches require a lot of effort to solve the desired parameterization due to the amount of vertex numbers. On the other hand, these techniques do not take advantage of the matrix structures to improve the efficiency of numerical computations. In fact, one can show that, after a suitable permutation,  $L_C$  is indeed a symmetric positive semi-definite skew-Hamiltonian matrix and *B* is a low-rank positive semi-definite matrix. Based on the special structures of  $L_C$  and *B*, we can compute the SCP by solving an associated eigenvalue problem with size only related to the number of boundary vertices so that the problem size is therefore much more smaller than those of the previous methods. This motivates us to study efficient and robust methods, through the particular matrix structures, for solving the GEP arising from the SCP.

#### 1.1 Contributions

The contribution of this work can be divided into three components: (1) *nonequivalence deflation*: a deflation technique to transform the zero eigenvalues of a GEP into the infinite ones while preserving all the other eigenvalues and associated eigenvectors; (2) *null-space free compression*: an approach of the model reduction to reduce a GEP to a small-scale standard eigenvalue problem (SEP) based on the low-rank property; (3) SHILA *algorithm*: a novel symmetric skew-Hamiltonian Isotropic Lanczos Algorithm for solving the symmetric skew-Hamiltonian eigenvalue problem that can precisely split the duplicate eigenvalues and improve the convergence rate. According to these three theoretical frameworks and numerical algorithms, we propose a novel, efficient, accurate and robust eigensolver for the SCP [31]. Related concepts and techniques can also provide fundamental tools for the issue of spectral mesh processing.

## 1.2 Notations and Overview

The following notations are frequently used throughout this paper. Other notations will be clearly defined whenever they are used.

- $-n_{v}$  denotes the number of vertices;  $n_{i}$  denotes the number of interior vertices as well as internal boundaries (if any), while  $n_{b}$  represents the number of (external) boundary vertices.
- Upper case letters indicate matrices and bold face letters denote vectors.
- $I_n$  denotes the  $n \times n$  identity matrix with the given size n.
- $-\mathbf{e}_{j}$  denotes the *j*th column of the identity matrix  $I_{n}$  with specified *n*.
- $-\mathbf{1}_n$  denotes an *n*-vector whose elements are all 1.

- We adopt **0** to denote the zero vectors and matrices of appropriate sizes.
- We use  $\cdot^{\top}$  to denote the transpose of vectors or matrices.

This paper is organized as follows. In Sects. 2 and 3, we review the background on discrete conformal maps and related works on the SCPs, respectively. With these basic concepts and terminologies, we propose a novel and efficient eigensolver to deal with the GEP arising from [31]. The theoretical frameworks and practical implementations of our eigensolver are specifically and successively introduced in Sects. 4 and 5. Numerical experiments and comparisons are presented in Sect. 6 and the concluding remarks are given in Sect. 7.

# 2 Discrete Conformal Maps

For a smooth map  $f : \mathcal{X} \to \mathcal{U}$ , we define the Dirichlet energy  $\mathcal{E}_D(f)$  and the area of the image of f,  $\mathcal{A}(f)$  by

$$\mathcal{E}_D(f) = \frac{1}{2} \int_{\mathcal{X}} |\nabla f|^2 d\sigma \text{ and } \mathcal{A}(f) = \int_{\mathcal{X}} \det(J_f) d\sigma,$$

respectively, where  $J_f$  is the Jacobian matrix of f and  $d\sigma$  is the area element of the surface. The conformal energy of f is the difference of  $\mathcal{E}_D(f)$  and  $\mathcal{A}(f)$ , defined by

$$\mathcal{E}_C(f) = \mathcal{E}_D(f) - \mathcal{A}(f). \tag{1}$$

Thanks to the relation  $\mathcal{E}_D(f) \ge \mathcal{A}(f)$  [6,33], we always have  $\mathcal{E}_C(f) \ge 0$  with the equality only when f is a conformal map.

By the discretization approach, we take  $\mathcal{X}$  and  $\mathcal{U}$  to be triangular meshes in  $\mathbb{R}^3$  and  $\mathbb{R}^2$ , respectively. Let  $\mathbf{f} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = [u_1, \dots, u_n, v_1, \dots, v_n]^{\top}$  represent a piecewise linear map from  $\mathcal{X}$  to  $\mathcal{U}$ . Then the discrete Dirichlet energy can be expressed as

$$E_D(\mathbf{f}) = \frac{1}{2} \sum_{e_{ij}} \frac{\cot \theta_{ij} + \cot \theta_{ji}}{2} \left[ (u_i - u_j)^2 + (v_i - v_j)^2 \right] = \frac{1}{2} \mathbf{f}^\top L_D \mathbf{f}, \qquad (2)$$

where  $\theta_{ij}$  and  $\theta_{ji}$  are the two corner angles opposite to the edge  $e_{ij}$  connecting vertices i, j on  $\mathcal{X}$ , and

$$L_D = \begin{bmatrix} K & \mathbf{0} \\ \mathbf{0} & K \end{bmatrix} \in \mathbb{R}^{2n_{\mathrm{v}} \times 2n_{\mathrm{v}}}$$
(3)

is the discrete Laplacian matrix with

$$K_{ij} = \begin{cases} -\sum_{\ell \in N(i)} K_{i\ell} & \text{if } j = i, \\ -\frac{1}{2} (\cot \theta_{ij} + \cot \theta_{ji}) & \text{if } j \in N(i), \\ 0 & \text{otherwise,} \end{cases}$$

in which N(i) denotes the set of all 1-ring neighboring vertices of vertex *i*. Moreover, we note that *K* is symmetric positive semi-definite [19] and  $K\mathbf{1}_{n_v} = \mathbf{0}$ . On the other hand, the area of the parameterization can be expressed as

$$\mathcal{A}(\mathbf{f}) = \frac{1}{2} \sum_{e_{ij} \in \partial \mathcal{U}} (u_i v_j - u_j v_i) = \frac{1}{2} \mathbf{f}^\top A \mathbf{f},$$
(4)

Deringer

where  $\partial \mathcal{U}$  is the set of boundary edges of  $\mathcal{U}$ , and

$$A = \begin{bmatrix} \mathbf{0} & -S \\ S & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{2n_{\mathrm{v}} \times 2n_{\mathrm{v}}}$$
(5)

\_

is a symmetric matrix with  $S^{\top} = -S$  and  $S\mathbf{1}_{n_{\nabla}} = \mathbf{0}$ . Specifically, if  $(u_i, v_i)$  and  $(u_j, v_j)$  are the endpoints of a boundary edge on  $\mathcal{U}$ , then

$$u_{i}v_{j} - u_{j}v_{i} = \begin{bmatrix} u_{i} & u_{j} & v_{i} & v_{j} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_{i} \\ u_{j} \\ \vdots \\ v_{i} \\ v_{j} \end{bmatrix}$$

and the matrix A in (5) is the assembly of the above  $2 \times 2$  elementary block matrices. As a result of (1)–(5), the discrete conformal energy can then be represented in a quadratic form

$$E_C(\mathbf{f}) = \frac{1}{2} \mathbf{f}^\top L_C \mathbf{f},\tag{6}$$

where

$$L_C = L_D - A = \begin{bmatrix} K & S \\ -S & K \end{bmatrix}$$
(7)

with *K* and *S* being symmetric positive semi-definite and skew-symmetric, respectively. In Marchandise et al. [28], derived the same matrix structure and property based on the finite element method for the minimization problem of the quadratic energy  $E_C$  in (6).

*Remark 1* The matrix  $L_C$  satisfies the following properties.

(i)  $L_C$  is symmetric and skew-Hamiltonian, i.e.,  $L_C^{\top} = L_C$  and  $(L_C J_{n_v})^{\top} = -(L_C J_{n_v})$  with

$$J_{n_{\nu}} = \begin{bmatrix} \mathbf{0} & I_{n_{\nu}} \\ -I_{n_{\nu}} & \mathbf{0} \end{bmatrix}.$$
 (8)

From [39], there is a symplectic orthogonal  $U \in \mathbb{R}^{2n_{\nabla} \times 2n_{\nabla}}$  with  $U^{\top} J_{n_{\nabla}} U = J_{n_{\nabla}}$  and  $U^{\top} U = I_{2n_{\nabla}}$  such that

$$U^{\top}L_{C}U = \begin{bmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \Lambda \end{bmatrix},$$

where  $\Lambda$  is an  $n_v \times n_v$  diagonal matrix. Since  $L_C$  is additionally, at least in theory, *positive semi-definite*, the above equality implies that the eigenvalues of  $L_C$  are nonnegative and double.

(ii) Set

$$\mathbb{I}_{2} = \begin{bmatrix} \mathbf{1}_{n_{\mathrm{v}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{n_{\mathrm{v}}} \end{bmatrix} \in \mathbb{R}^{2n_{\mathrm{v}} \times 2},\tag{9}$$

where  $\mathbf{1}_{n_{v}}$  be an  $n_{v}$ -vector as defined in Sect. 1.2. Then we have

$$L_C \mathbb{I}_2 = \mathbf{0} \tag{10}$$

as  $K\mathbf{1}_{n_v} = \mathbf{0} = S\mathbf{1}_{n_v}$ . Moreover, we conclude that 0 is a semisimple eigenvalue of  $L_C$  with multiplicity 2.

#### **3** The Spectral Conformal Parameterization

We briefly review the connection between the minimization problem of the conformal energy (6) and the associated GEP in SCP [31].

#### 3.1 Minimization of Conformal Energy

The piecewise linear mapping **f** is called a DCP if it minimizes the conformal energy  $E_C(\mathbf{f})$  in (6). Lévy et al. [25] and Desbrun et al. [6] independently proposed LSCM and DCP, which are theoretically equivalent, to achieve such an energy minimization problem. Through pinning down two vertices in the parameter region to eliminate the inherent rank deficiency, the non-trivial parameterization result ( $\mathbf{f} \neq \text{constant}$ ) of LSCM/DCP can be uniquely determined by solving the linear system  $L_C \mathbf{f} = \mathbf{0}$ . The choice of which vertices to be fixed, however, significantly affects the quality of the conformal parameterization.

To remedy this problem, Mullen et al. [31] suggested a spectral approach based on the so-called (generalized) Fiedler vector [7] to avoid the explicit constraint on specific vertices. To this end, Mullen et al. focused on the following constrained minimization problem:

$$\min_{\mathbf{f}} \mathbf{f}^{\top} L_C \mathbf{f} \quad \text{subject to} \quad \mathbf{f}^{\top} B \mathbb{I}_2 = \mathbf{0}, \ \mathbf{f}^{\top} B \mathbf{f} = 1,$$
(11)

where  $E_C$  and  $\mathbb{I}_2$  are defined as in (6) and (9), respectively, and *B* is a degenerate and diagonal binary matrix whose nonzero elements correspond to the (external) boundary vertices. Note that *B* can be expressed as the block-diagonal form

$$B = \begin{bmatrix} D & \mathbf{0} \\ \mathbf{0} & D \end{bmatrix} \in \mathbb{R}^{2n_{\mathrm{v}} \times 2n_{\mathrm{v}}},\tag{12}$$

where *D* is an  $n_v \times n_v$  diagonal binary matrix with 1 at the diagonal entries corresponding to the boundary vertices (not including any of internal boundaries). The constraints in (11) indicate that the barycenter of the boundary components must be at zero ( $\mathbf{f}^\top B \mathbb{I}_2 = \mathbf{0}$ ), and the moment of inertia on the boundary must be unit ( $\mathbf{f}^\top B \mathbf{f} = 1$ ).

The following lemma shows that, solving the optimization problem (11) is equivalent to finding the eigenvector of the GEP

$$L_C \mathbf{f} = \lambda B \mathbf{f} \tag{13}$$

corresponding to the smallest positive eigenvalue.

**Lemma 1** A vector  $\mathbf{f}_*$  is an optimizer of the constrained energy minimization problem (11) if and only if  $\mathbf{f}_*$  is the eigenvector of the GEP (13) corresponding to the smallest positive eigenvalue with  $\mathbf{f}_*^\top B \mathbf{f}_* = 1$ .

Proof Consider the Lagrange function

$$\mathcal{L}(\mathbf{f},\mu,\lambda) = \mathbf{f}^{\top} L_C \mathbf{f} - \mu \left( \mathbf{f}^{\top} B \mathbb{I}_2 \right) - \lambda \left( \mathbf{f}^{\top} B \mathbf{f} - 1 \right)$$

with Lagrange multipliers  $\mu$  and  $\lambda$ . Then the solution **f** satisfies

$$\frac{\partial \mathcal{L}}{\partial \mathbf{f}} = 2L_C \mathbf{f} - \mu B \mathbb{I}_2 - 2\lambda B \mathbf{f} = \mathbf{0}, \tag{14a}$$

$$\frac{\partial \mathcal{L}}{\partial \mu} = \mathbf{f}^{\top} B \mathbb{I}_2 = \mathbf{0}, \tag{14b}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \mathbf{f}^{\top} B \mathbf{f} = 1.$$
(14c)

Deringer

Premultiplying (14a) by  $\mathbb{I}_2^{\top}$ , we get

$$\mu(n_{\mathrm{b}}I_2) = 2 \left( L_C \mathbb{I}_2 \right)^\top \mathbf{f} - 2\lambda \left( \mathbf{f}^\top B \mathbb{I}_2 \right)^\top = \mathbf{0}, \tag{15}$$

where the last equality holds because of (10) and (14b). From (15), we must have  $\mu = 0$ . As a result, (14a) can be reduced to the GEP (13).

Conversely, if  $\mathbf{f}_*$  is the eigenvector corresponding to the smallest positive eigenvalue  $\lambda_*$  of the GEP (13) with the normalizing requirement  $\mathbf{f}_*^\top B \mathbf{f}_* = 1$ . Then, by (10), we see that

$$\mathbf{0} = (L_C \mathbb{I}_2)^\top \mathbf{f}_* = \mathbb{I}_2^\top (L_C \mathbf{f}_*) = \mathbb{I}_2^\top (\lambda_* B \mathbf{f}_*), \qquad (16)$$

which implies that  $\mathbf{f}_*^{\top} B \mathbb{I}_2 = \mathbf{0}$ . So,  $\mathbf{f}_*$  satisfies the requirements of constraints in (11). Furthermore, for any vector  $\mathbf{g}$  satisfying  $\mathbf{g}^{\top} B \mathbb{I}_2 = \mathbf{0}$  and  $\mathbf{g}^{\top} B \mathbf{g} = 1$ , we have

$$\mathbf{g}^{\top} L_C \mathbf{g} \geq \min_{\mathbf{f}} \left\{ \mathbf{f}^{\top} L_C \mathbf{f} : \mathbf{f}^{\top} B \mathbb{I}_2 = \mathbf{0}, \ \mathbf{f}^{\top} B \mathbf{f} = 1 \right\} = \lambda_* = \mathbf{f}_*^{\top} L_C \mathbf{f}_*$$

Thus,  $\mathbf{f}_*$  solve the constrained minimization problem (11).

In other words, the solution to (13) with the smallest positive eigenvalue determines the entire coordinates of the SCP, named by Mullen et al. [31].

# 3.2 Previous Numerical Methods

To treat the GEP (13), Mullen et al. [31] considered the modified GEP:

$$\left[B - \frac{1}{n_{\rm b}} (B\mathbb{I}_2) (B\mathbb{I}_2)^{\top}\right] \mathbf{f} = \frac{1}{\lambda} L_C \mathbf{f}.$$
(17)

By taking  $\mathbf{f} = \begin{bmatrix} \mathbf{1}_{n_v} \\ \mathbf{0} \end{bmatrix}$  or  $\mathbf{f} = \begin{bmatrix} \mathbf{0} \\ \mathbf{1}_{n_v} \end{bmatrix}$ , we see that both sides of (17) are equal to zero vectors which means that it is a *singular* GEP. So, the modified GEP (17) can be ill posed in the sense that an arbitrary small perturbation may cause a large change of the eigenvalues and the associated eigenvectors [3,30]. Later, we will see that although  $(B\mathbb{I}_2)^{\top}\mathbf{f}$  is theoretically equivalent to zero, it, in practice, still has a significant numerical error and can further affect the residual norm  $\|L_C \mathbf{f} - \lambda B\mathbf{f}\|_2$ .

Alexa and Wardetzky [1] recently addressed the GEP (13) via an equivalent smaller eigenvalue problem that only contains the boundary vertices. By reordering the vertex indices, the GEP (13) can be rewritten as

$$\begin{bmatrix} L_{ii} & L_{ib} \\ L_{ib}^{\top} & L_{bb} \end{bmatrix} \begin{bmatrix} \mathbf{f}_i \\ \mathbf{f}_b \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & B_b \end{bmatrix} \begin{bmatrix} \mathbf{f}_i \\ \mathbf{f}_b \end{bmatrix},$$
(18)

where i and b refer as to inner and (external) boundary vertices, respectively,  $\mathbf{f}_{i} = \begin{bmatrix} \mathbf{u}_{i} \\ \mathbf{v}_{i} \end{bmatrix} \in \mathbf{F}_{i}$ 

 $\mathbb{R}^{2n_{\pm}}$ ,  $\mathbf{f}_{\mathrm{b}} = \begin{bmatrix} \mathbf{u}_{\mathrm{b}} \\ \mathbf{v}_{\mathrm{b}} \end{bmatrix} \in \mathbb{R}^{2n_{\mathrm{b}}}$ , and  $B_{\mathrm{b}}$  is a  $2n_{\mathrm{b}} \times 2n_{\mathrm{b}}$  diagonal matrix with 1's on the diagonal cor-

responding to the (external) boundary vertices.<sup>1</sup> Under such a structure, Alexa and Wardetzky considered the Schur complement of  $L_{ii}$  from (18), i.e.,  $L_{bb} - L_{ib}^{\top} L_{ii}^{-1} L_{ib} \in \mathbb{R}^{2n_b \times 2n_b}$ . Thus, (18) can be reduced to a small-scale GEP

$$\left(L_{\rm bb} - L_{\rm ib}^{\top} L_{\rm ii}^{-1} L_{\rm ib}\right) \mathbf{f}_{\rm b} = \lambda B_{\rm b} \mathbf{f}_{\rm b},\tag{19}$$

<sup>&</sup>lt;sup>1</sup> For mesh with single boundary, we have  $B_{\rm b} = I_{2n_{\rm b}}$ .

and  $\mathbf{f}_{i}$  as well as  $\mathbf{f}_{b}$  satisfy the relationship

$$\mathbf{f}_{i} = -L_{ij}^{-1}L_{ib}\mathbf{f}_{b}.$$
(20)

Compared with Mullen et al. who directly cope with the GEP (13), Alexa and Wardetzky first handle the reduced GEP (19) to obtain  $\mathbf{f}_{\text{D}}$  and then determine the parameterization coordinates of interior vertices  $\mathbf{f}_{\pm}$  via the Eq. (20).

Yet this approach still has some drawbacks and difficulties. First of all,  $\lambda = 0$  remains a semisimple eigenvalue of the reduced GEP (19) associated with the eigenvectors  $\mathbf{f}_{\text{D}} = \begin{bmatrix} \mathbf{1}_{n_{\text{D}}} \\ \mathbf{0} \end{bmatrix}$ 

and  $\mathbf{f}_{\rm b} = \begin{bmatrix} \mathbf{0} \\ \mathbf{1}_{n_{\rm b}} \end{bmatrix}$  so its kernel still affects the convergence and increases the computational cost. Secondly, to deal with the reduced GEP (19), we may first face the problem of solving a linear system

$$L_{ii}X = L_{ib}.$$
 (21)

Even though  $L_{\pm\pm}$  is well-conditioned, it can be impractical or even impossible to solve the linear system beforehand just in order to solve "one" desired eigenvector. Here, the main reason is that the number of right hand sides equals the number of all boundary vertices which may be more than hundreds or thousands of points, in particular for large meshes. Last but not least, as opposed to solving the linear system (21) in advance, we may adopt the iterative method to solve (19). The computational cost is due to the problem of inner-outer iterations. Especially for the inner step, we need to solve a singular linear system with the coefficient matrix  $L_{\rm bb} - L_{\rm ib}^{\top} L_{\rm ib}^{-1} L_{\rm ib}$ .

### 4 Theoretical Frameworks

In this section, we will propose a clever compression technique together with a novel eigensolver for solving the smallest positive eigenvalue and associated eigenvector of the GEP (13)

$$L_C \mathbf{f} = \lambda B \mathbf{f},$$

where  $L_C$  is symmetric positive semi-definite, skew-Hamiltonian as defined in (7) and *B*, given in (12), is symmetric positive semi-definite.

#### 4.1 Nonequivalence Deflations

The GEP (13) possesses the zero eigenvalue with algebraic multiplicity 2 and, in fact,  $\operatorname{Ker}(L_C) = \operatorname{span}\left\{\begin{bmatrix}\mathbf{1}_{n_v}\\\mathbf{0}\end{bmatrix}, \begin{bmatrix}\mathbf{0}\\\mathbf{1}_{n_v}\end{bmatrix}\right\}$ . To find the eigenpairs associated with the smallest positive eigenvalue of the GEP (13) and, at the same time, to exclude undesired eigenpairs corresponding the zero eigenvalue, we introduce a *nonequivalence deflation* technique, based on the idea in [26,27], to transform the zero eigenvalues of the GEP (13) into the infinite ones while preserving all the other eigenvalues of (13).

Observe that since

$$\mathbf{d} \equiv D\mathbf{1}_{n_{\mathrm{v}}} \neq \mathbf{0},\tag{22}$$

we have  $B\mathbb{I}_2 = \begin{bmatrix} \mathbf{d} & \mathbf{0} \\ \mathbf{0} & \mathbf{d} \end{bmatrix} \neq \mathbf{0}$ , where D and  $\mathbb{I}_2$  are defined in (12) and (9), respectively. The *nonequivalence transformation* of the matrix pair  $(L_C, B)$  is defined by

Deringer

$$\widetilde{L}_C \equiv L_C + \left(\frac{1}{n_{\rm b}}B\mathbb{I}_2\right)(B\mathbb{I}_2)^{\top} = \begin{bmatrix} \widetilde{K}S\\ -S \ \widetilde{K} \end{bmatrix}, \quad \widetilde{K} = K + \frac{1}{n_{\rm b}}\mathbf{d}\mathbf{d}^{\top}, \quad (23a)$$

$$\widetilde{B} \equiv B - \left(\frac{1}{n_{\rm b}}B\mathbb{I}_2\right)(B\mathbb{I}_2)^{\top} = \begin{bmatrix} \widetilde{D} & \mathbf{0} \\ \mathbf{0} & \widetilde{D} \end{bmatrix}, \quad \widetilde{D} = D - \frac{1}{n_{\rm b}}\mathbf{d}\mathbf{d}^{\top}.$$
(23b)

Observe that

$$L_C \mathbb{I}_2 = B \mathbb{I}_2 \neq \mathbf{0},\tag{24}$$

and

$$\widetilde{B}\mathbb{I}_2 = \mathbf{0}.\tag{25}$$

This indicates that such a nonequivalence deflation technique transforms the kernel matrix  $\mathbb{I}_2$  of  $L_C$  to be parts of the kernel of  $\tilde{B}$ . As a matter of fact, this technique can tell us more about the spectral behavior.

To begin with, we first investigate the spectrum of the deflated matrix  $\tilde{D}$  in (23b). Let  $\mathcal{I}_{\pm}$  and  $\mathcal{I}_{b} = \{b_1, \ldots, b_{n_b}\}$  denote *ordered* index sets of the interior vertices (including interior boundary vertices) and the (external) boundary vertices, respectively. Let  $\mathbf{b}_k$  be an  $n_{\nabla}$ -vector defined by

$$(\mathbf{b}_k)_i = \begin{cases} \frac{1}{\sqrt{k(k+1)}}, \ i = b_1, \cdots, b_k, \\ \frac{-k}{\sqrt{k(k+1)}}, \ i = b_{k+1}, \\ 0, & \text{otherwise}, \end{cases} \quad k = 1, \dots, n_{\mathrm{b}} - 1.$$
(26)

**Lemma 2** The deflated matrix  $\widetilde{D}$  in (23b) has semisimple eigenvalues 0 and 1 with algebraic multiplicity  $n_{i} + 1$  and  $n_{b} - 1$ , respectively. Moreover, the kernel of  $\widetilde{D}$  has an orthonormal basis  $\left\{ \mathbf{e}_{j}, \frac{1}{\sqrt{n_{b}}} \mathbf{d} : j \in \mathcal{I}_{i} \right\}$ , where  $\mathbf{d}$  is defined in (22); the eigenspace of  $\widetilde{D}$  corresponding to the eigenvalue 1 is spanned by the orthonormal set  $\left\{ \mathbf{b}_{1}, \ldots, \mathbf{b}_{n_{b}-1} \right\}$ .

*Proof* For simplicity, we first reorder the columns and rows of D in (12) to get

$$D = \frac{n_{\rm b}}{n_{\rm i}} \begin{bmatrix} n_{\rm b} & n_{\rm i} \\ I_{n_{\rm b}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Then, we have

$$\widetilde{D} = D - \frac{1}{n_{\rm b}} \mathbf{d} \mathbf{d}^{\top} = \begin{bmatrix} I_{n_{\rm b}} - \frac{1}{n_{\rm b}} \mathbf{1}_{n_{\rm b}} \mathbf{1}_{n_{\rm b}}^{\top} \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{d} = D \mathbf{1}_n = \begin{bmatrix} \mathbf{1}_{n_{\rm b}} \\ \mathbf{0} \end{bmatrix},$$
$$\mathcal{I}_{\rm b} = \{1, \dots, n_{\rm b}\}, \quad \mathcal{I}_{\rm b} = \{n_{\rm b} + 1, \dots, n\},$$

and (26) can be rewritten as

$$\mathbf{b}_{k} = \frac{1}{\sqrt{k(k+1)}} \begin{bmatrix} \mathbf{1}_{k} \\ -k \\ \mathbf{0} \end{bmatrix}, \quad k = 1, \dots, n_{\mathrm{b}} - 1.$$
(27)

The orthogonality of the  $n_{v}$  vectors  $\{\mathbf{b}_{1}, \ldots, \mathbf{b}_{n_{\mathrm{b}}-1}, \frac{1}{\sqrt{n_{\mathrm{b}}}}\mathbf{d}, \mathbf{e}_{n_{\mathrm{b}}+1}, \ldots, \mathbf{e}_{n_{v}}\}$  is straightforward via a simple calculation. It is obvious to see that  $\widetilde{D}\mathbf{d} = \mathbf{0}$  and  $\widetilde{D}\mathbf{e}_{k} = \mathbf{0}$  for each  $k \in \mathcal{I}_{1}$  which imply that the  $n_{1} + 1$  vectors,  $\frac{1}{\sqrt{n_{\mathrm{b}}}}\mathbf{d}, \mathbf{e}_{n_{\mathrm{b}}+1}, \ldots, \mathbf{e}_{n_{v}}$ , are eigenvectors of  $\widetilde{D}$  corresponding to the zero eigenvalue. On the other hand, owing to the orthogonality of  $\mathbf{d}^{\mathsf{T}}\mathbf{b}_{k} = 0$  for  $k = 1, \ldots, n_{\mathrm{b}} - 1$ , we have  $\widetilde{D}\mathbf{b}_{k} = D\mathbf{b}_{k} = \mathbf{b}_{k}$ . This shows that the  $n_{\mathrm{b}} - 1$  vectors,  $\mathbf{b}_{1}, \ldots, \mathbf{b}_{n_{\mathrm{b}}-1}$ , are eigenvectors of  $\widetilde{D}$  corresponding to the eigenvalue 1.

**Theorem 1** Let  $(\tilde{L}_C, \tilde{B})$  be the deflated pair as defined in (23). Then

- (i)  $\tilde{L}_C$  is symmetric positive definite and skew-Hamiltonian.
- (ii)  $\tilde{B}$  is symmetric positive semi-definite with  $2(n_{\pm} + 1)$ 's semisimple eigenvalues 0 and  $2(n_{b} 1)$ 's semisimple eigenvalues 1.
- (*iii*) The deflated GEP

$$\widetilde{L}_C \mathbf{f} = \lambda \widetilde{B} \mathbf{f} \tag{28}$$

preserves all eigenpairs of the original GEP (13) except the case that  $\lambda = 0$ . Instead, two semisimple zero eigenvalues of (13) are transformed into two semisimple infinite

eigenvalues of (28) associated with eigenvectors  $\begin{bmatrix} \mathbf{1}_{n_{v}} \\ \mathbf{0} \end{bmatrix}$  and  $\begin{bmatrix} \mathbf{0} \\ \mathbf{1}_{n_{v}} \end{bmatrix}$ .

*Proof* (i) Clearly,  $\tilde{L}_C$  is symmetric and positive definite directly from the facts that  $L_C$  as well as  $\frac{1}{n_{\text{E}}}(B\mathbb{I}_2)(B\mathbb{I}_2)^{\top}$  are both symmetric positive semi-definite, and their individual Rayleigh quotient cannot vanish simultaneously as

$$\operatorname{Ker}(L_C) \cap \operatorname{Ker}((B\mathbb{I}_2)(B\mathbb{I}_2)^{\perp}) = \emptyset.$$

Let  $J_{n_{v}}$  be the matrix in (8). It is easy to verify that  $J_{n_{v}}B = BJ_{n_{v}}$  and  $J_{n_{v}}\mathbb{I}_{2}\mathbb{I}_{2}^{\top} = \mathbb{I}_{2}\mathbb{I}_{2}^{\top}J_{n_{v}}$ . Therefore, noting that  $L_{C}$  is a skew-Hamiltonian matrix and  $J_{n_{v}}^{\top} = -J_{n_{v}}$ , we get

$$(\widetilde{L}_C J_{n_{\mathrm{v}}})^{\top} = -L_C J_{n_{\mathrm{v}}} - \frac{1}{n_{\mathrm{b}}} (B\mathbb{I}_2) (B\mathbb{I}_2)^{\top} J_{n_{\mathrm{v}}} = -(\widetilde{L}_C J_{n_{\mathrm{v}}}),$$

i.e.,  $\tilde{L}_C$  is also a skew-Hamiltonian matrix.

(ii) Since  $\tilde{B}$  is a block-diagonal matrix composed of the deflated matrix  $\tilde{D}$  as in (23b), from Lemma 2, we conclude that the nullity of  $\tilde{B}$  is equal to  $2(n_{\pm} + 1)$  and the dimension of the eigenspace associated with the eigenvalue 1 of  $\tilde{B}$  is  $2(n_{\rm b} - 1)$ .

(iii) By the same proof technique as used for (16), we first observe that if  $(\theta, \mathbf{g})$  is an eigenpair of the GEP (13) with  $\theta \neq 0$  then,  $\mathbb{I}_2$  and  $\mathbf{g}$  are *B*-orthogonal, i.e.,  $\mathbb{I}_2^\top B \mathbf{g} = \mathbf{0}$ . As a result, from (23a) and (23b), we see that

$$\widetilde{L}_C \mathbf{g} = L_C \mathbf{g} = \theta B \mathbf{g} = \theta \widetilde{B} \mathbf{g}.$$
(29)

In addition, (24) and (25) imply that  $\begin{bmatrix} \mathbf{1}_{n_{v}} \\ \mathbf{0} \end{bmatrix}$  and  $\begin{bmatrix} \mathbf{0} \\ \mathbf{1}_{n_{v}} \end{bmatrix}$  are eigenvectors of the deflated GEP (28) corresponding to the infinite eigenvalues.

It is well known that the iterative projection methods, such as the Lanczos method or the Arnoldi method, rapidly provide approximate eigenvalues with large magnitude. To compute the smallest positive eigenvalue(s) and the associated eigenvector(s) of the deflated GEP (28), we may invert the  $\tilde{L}_C$  and straightforwardly apply the Lanczos method to the SEP of the form

$$(\tilde{L}_C^{-1}\tilde{B})\mathbf{f} = \frac{1}{\lambda}\mathbf{f}.$$
(30)

However, to deal directly with the problem (30) does not make use of the special structures of the coefficient matrices that  $\tilde{L}_C$  is symmetric positive definite, skew-Hamiltonian and  $\tilde{B}$  is symmetric positive semi-definite with low-rank. In the subsequent subsections, we will first expound how to draw on the low-rank property of  $\tilde{B}$  to reduce the SEP (30) to a symmetric positive definite and skew-Hamiltonian eigenvalue problem which is of the small size  $2(n_{\rm b} - 1)$ . Then, we will propose a modified Lanczos algorithm to solve this reduced problem efficiently.

## 4.2 Null-Space Free Compression

The matrix B is a diagonal matrix with rank  $2n_{\rm b}$  and the subsequent rank-two update deflating procedure causes the rank of  $\tilde{B}$  to be deficient by two (see Theorem 1 and Lemma 2). Compared with the matrix size of  $\tilde{B}$ ,  $2n_{\rm v}$ , the rank of  $\tilde{B}$ ,  $2(n_{\rm b} - 1)$ , seems "extremely small". Therefore, a low-rank compression technique on  $\tilde{B}$  will have the benefits of reducing the matrix size, computational cost and memory storage for efficiently solving the SEP (30). As we mentioned above in Theorem 1, 0 and 1 are the only two eigenvalues of  $\tilde{B}$  and, particularly, all of its eigenvectors can be formulated explicitly. We will also return to this point in Sect. 5.1.

We next explain the concept of low-rank compression to reduce the SEP (30).

**Lemma 3** Let  $\tilde{B}_1$  be a  $2n_v \times 2(n_b - 1)$  column orthonormal matrix whose columns form an orthonormal basis of the eigenspace of  $\tilde{B}$  associated with the eigenvalue 1. Then  $\tilde{B}_1$  can be represented in the form:

$$\widetilde{B}_1 = \begin{bmatrix} E_1 & \mathbf{0} \\ \mathbf{0} & E_1 \end{bmatrix},\tag{31}$$

where  $E_1 \in \mathbb{R}^{n_v \times (n_b - 1)}$  satisfying  $E_1 E_1^\top = \widetilde{D}$  with  $\widetilde{D}$  as in (23b).

*Proof* By Lemma 2, it holds that  $\operatorname{rank}(\widetilde{D}) = n_{\rm b} - 1$ . Let  $\widetilde{D} = E_1 E_1^{\top}$  be a low-rank compression of  $\widetilde{D}$ . Then, according to the matrix structure of  $\widetilde{B}$  in (23b), we see that

$$\widetilde{B} = \begin{bmatrix} \widetilde{D} & \mathbf{0} \\ \mathbf{0} & \widetilde{D} \end{bmatrix} = \begin{bmatrix} E_1 E_1^\top & \mathbf{0} \\ \mathbf{0} & E_1 E_1^\top \end{bmatrix} = \begin{bmatrix} E_1 & \mathbf{0} \\ \mathbf{0} & E_1 \end{bmatrix} \begin{bmatrix} E_1^\top & \mathbf{0} \\ \mathbf{0} & E_1^\top \end{bmatrix} \equiv \widetilde{B}_1 \widetilde{B}_1^\top.$$

**Theorem 2** Under the assumption in Lemma 3, the SEP (30) can be reduced to the smallscale, compressed and deflated SEP

$$(\widetilde{B}_1^{\top} \widetilde{L}_C^{-1} \widetilde{B}_1) \mathbf{s}_1 = \frac{1}{\lambda} \mathbf{s}_1, \tag{32}$$

where  $\widetilde{B}_1^{\top} \widetilde{L}_C^{-1} \widetilde{B}_1$  is symmetric positive definite and skew-Hamiltonian with size  $2(n_{\rm b} - 1)$ . From now on, we use CDSEP to indicate the eigenvalue problem (32).

*Proof* Since  $\hat{B}$  is symmetric positive semi-definite, we can first rewrite the matrix  $\hat{B}$  to a condensed form

$$\widetilde{B} = \widetilde{B}_1 \widetilde{B}_1^\top, \tag{33}$$

where  $\tilde{B}_1$  is a  $2n_v \times 2(n_b - 1)$  matrix as in (31). Moreover, if  $\tilde{B}_0$  is a  $2n_v \times 2(n_{\pm} + 1)$  orthonormal matrix so that its columns span the kernel of  $\tilde{B}$ , then any  $2n_v$ -vector **f** can be uniquely expressed as a linear combination of  $\tilde{B}_0$  and  $\tilde{B}_1$ , that is,

$$\mathbf{f} = \widetilde{B}_0 \mathbf{s}_0 + \widetilde{B}_1 \mathbf{s}_1 \tag{34}$$

for some  $\mathbf{s}_0 \in \mathbb{R}^{2(n_{\rm b}+1)}$  and  $\mathbf{s}_1 \in \mathbb{R}^{2(n_{\rm b}-1)}$ . Substituting Eqs. (33) and (34) into (30), and premultiplying the resulting equation by  $\widetilde{B}_1^{\top}$ , we can further reduce the SEP (30) to the small-scale CDSEP (32) with size  $2(n_{\rm b}-1)$  due to the orthogonality  $\widetilde{B}_1^{\top} \widetilde{B}_0 = \mathbf{0}$ .

It is obviously that  $\widetilde{B}_1^{\top} \widetilde{L}_C^{-1} \widetilde{B}_1$  is symmetric. Since  $\widetilde{B}_1$  is column orthonormal, by Theorem 1 (i), we conclude that  $\widetilde{B}_1^{\top} \widetilde{L}_C^{-1} \widetilde{B}_1$  is positive definite. We now show that  $\widetilde{B}_1^{\top} \widetilde{L}_C^{-1} \widetilde{B}_1$  is skew-Hamiltonian. Let  $J_{\ell} = \begin{bmatrix} \mathbf{0} & -I_{\ell} \\ I_{\ell} & \mathbf{0} \end{bmatrix}$  with  $\ell = 2(n_{\rm b} - 1)$ . According to the block-diagonal-like form of  $\widetilde{B}_1$  in (31), we see that  $\widetilde{B}_1 J_\ell = J_{n_\nu} \widetilde{B}_1$  where  $J_{n_\nu}$  is the matrix defined in (8). Since  $J_{n_\nu}^\top = -J_{n_\nu} = J_{n_\nu}^{-1}$  and  $\widetilde{L}_C$  is symmetric skew-Hamiltonian, it is easy to verify that  $\widetilde{L}_C^{-1} J_{n_\nu} = J_{n_\nu} \widetilde{L}_C^{-1}$ . Therefore, we can deduce that

$$(\widetilde{B}_1^{\top} \widetilde{L}_C^{-1} \widetilde{B}_1 J_\ell)^{\top} = (J_\ell \widetilde{B}_1^{\top} \widetilde{L}_C^{-1} \widetilde{B}_1)^{\top} = -\widetilde{B}_1^{\top} \widetilde{L}_C^{-1} \widetilde{B}_1 J_\ell.$$

From (30)–(34), as soon as we obtain an eigenpair ( $\lambda^{-1}$ ,  $\mathbf{s}_1$ ) of the CDSEP (32), the desired eigenvector **f** of the deflated GEP (28) (and hence of the original GEP (13)) can be obtained directly through the relation

$$\mathbf{f} = \lambda \widetilde{L}_C^{-1} \widetilde{B}_1 \mathbf{s}_1. \tag{35}$$

Based on (16) in Lemma 1, **f** and  $\mathbb{I}_2$  must be *B*-orthogonal. In addition, from (32) and (35), since

$$\widetilde{B}_1^{\top}\mathbf{f} = \lambda(\widetilde{B}_1^{\top}\widetilde{L}_C^{-1}\widetilde{B}_1)\mathbf{s}_1 = \mathbf{s}_1,$$

the *B*-orthogonality of  $\mathbf{f}$  and  $\mathbb{I}_2$  implies that

$$\mathbf{f}^{\top} B \mathbf{f} = \mathbf{f}^{\top} \left( \widetilde{B} + \left( \frac{1}{n_{\mathrm{b}}} B \mathbb{I}_2 \right) (B \mathbb{I}_2)^{\top} \right) \mathbf{f} = \mathbf{f}^{\top} \widetilde{B}_1 \widetilde{B}_1^{\top} \mathbf{f} = \mathbf{s}_1^{\top} \mathbf{s}_1$$

Consequently,  $\mathbf{f}^{\top} B \mathbf{f} = 1$  provided that  $\mathbf{s}_1^{\top} \mathbf{s}_1 = 1$ .

*Remark* 2 We remark on the difference between these two eigenproblems (19) and (32). In the first place, the reduced problem (32) excludes the possibility of interference induced by the kernel of  $L_C$  [Theorem 1 (i)]. Second, the inverted CDSEP (32) allows us to seek the smallest positive eigenvalue directly through any well-known iterative methods without the need to previously construct the coefficient matrix. To put it another way, we only need to devise effective techniques to compute the matrix-vector multiplication  $\tilde{B}_1\mathbf{q}$ , and to solve the linear system  $\tilde{L}_C \mathbf{p} = \mathbf{r}$  for given vectors  $\mathbf{q}$ ,  $\mathbf{r}$ . Thus, for large meshes, through the CDSEP (32) to determine the smallest positive eigenvalue and its associated eigenvector of the original GEP (13) is more efficient and robust than the equivalent problem (19).

4.3 SHILA: The Symmetric Skew-Hamiltonian Isotropic Lanczos Algorithm

Subsequently, we will propose a novel and efficient eigensolver for the SEP  $Ms = \lambda s$  with a symmetric skew-Hamiltonian operator M. In our case, M is given by  $\tilde{B}_1^{\top} \tilde{L}_C^{-1} \tilde{B}_1$  and the practical realization will be discussed in Sect. 5.

Let *n* be a positive integer. Suppose that  $M \in \mathbb{R}^{2n \times 2n}$  is a skew-Hamiltonian matrix (not necessarily symmetric). Van Loan [40], showed that there is a  $2n \times 2n$  symplectic<sup>2</sup> and orthogonal matrix of the form  $\begin{bmatrix} Q & JQ \end{bmatrix}$  with  $Q \in \mathbb{R}^{2n \times n}$  and  $J = \begin{bmatrix} \mathbf{0} & -I_n \\ I_n & \mathbf{0} \end{bmatrix}$  such that

$$\begin{bmatrix} Q^{\top} \\ (JQ)^{\top} \end{bmatrix} M \begin{bmatrix} Q & JQ \end{bmatrix} = \begin{bmatrix} H & F \\ \mathbf{0} & H^{\top} \end{bmatrix},$$
(36)

where  $H \in \mathbb{R}^{n \times n}$  is upper Hessenberg and  $F \in \mathbb{R}^{n \times n}$  is skew-symmetric. The matrix structure in (36) presents the multiplicity of the eigenvalues of M. For each double eigenvalue of M, one copy resides in H, and the other copy is in  $H^{\top}$ . Consequently, the eigenvalues of M can be captured by H without missing any information.

<sup>&</sup>lt;sup>2</sup> A  $2n \times 2n$  matrix G is said to be symplectic if  $G^{\top}JG = J$ .

Recall that the *Krylov subspace*  $\mathcal{K}_k(M, \mathbf{q})$  of *M* with respect to  $\mathbf{q}$  and *k* is defined by

$$\mathcal{K}_k(M, \mathbf{q}) = \operatorname{span}\left\{\mathbf{q}, M\mathbf{q}, M^2\mathbf{q}, \dots, M^{k-1}\mathbf{q}\right\}.$$

It can be shown that when M is skew-Hamiltonian, the associated Krylov subspace is *isotropic*, which means that  $\mathbf{s}^{\top} J \mathbf{t} = 0$  for all  $\mathbf{s}, \mathbf{t} \in \mathcal{K}_k(M, \mathbf{q})$  [29]. To compute an orthonormal and symplectic basis  $\{\mathbf{q}_j\}_{j=1}^k$  of such a *k*-dimensional isotropic Krylov subspace, Mehrmann and Watkins [29] introduced the so-called *isotropic Arnoldi process* 

$$\mathbf{q}_{j+1}h_{j+1,j} = M\mathbf{q}_j - \sum_{i=1}^j \mathbf{q}_i h_{ij} - \sum_{i=1}^j J\mathbf{q}_i r_{ij}, \quad j = 1, \dots, k-1,$$
(37)

where  $h_{ij} = \mathbf{q}_i^\top M \mathbf{q}_j$ ,  $h_{j+1,j}$  is chosen to be a positive number so that  $\|\mathbf{q}_{j+1}\|_2 = 1$  and

$$r_{ij} = (J\mathbf{q}_i)^\top M\mathbf{q}_j. \tag{38}$$

For a skew-Hamiltonian matrix M with exact arithmetic,  $r_{ij}$  in (37) will all be zero, so the isotropic Arnoldi process and the standard Arnoldi process are theoretically equivalent. However, in the practical implementation, some tiny values for  $r_{ij}$  caused by roundoff error will destroy the isotropic property. To prevent the loss of isotropicity, we need to subtract out the tiny component  $J\mathbf{q}_i r_{ij}$ . The process terminates after n - 1 steps as  $\{\mathbf{q}_1, \ldots, \mathbf{q}_n, J\mathbf{q}_1, \ldots, J\mathbf{q}_n\}$  forms an orthonormal basis of  $\mathbb{R}^{2n}$ . Based on the isotropic Arnoldi process, Mehrmann and Watkins [29] further developed the SHIRA method, which is the abbreviation of *skew-Hamiltonian, isotropic, implicitly restarted shift-and-invert Arnoldi method*, for solving for the large-scale SEP  $M\mathbf{s} = \lambda \mathbf{s}$  with a real skew-Hamiltonian operator M.

For our model problem,  $M = \tilde{B}_1^{\top} \tilde{L}_C^{-1} \tilde{B}_1$  in (32) is skew-Hamiltonian, and, additionally, itself is *symmetric*. In this case, the equality in (36) can be reduced as

$$\begin{bmatrix} Q^{\top} \\ (JQ)^{\top} \end{bmatrix} M \begin{bmatrix} Q & JQ \end{bmatrix} = \begin{bmatrix} T & \mathbf{0} \\ \mathbf{0} & T \end{bmatrix},$$

where T is an  $n \times n$  tridiagonal matrix. Based on the isotropic Arnoldi process, we now introduce an *isotropic Lanczos procedure* for a symmetric skew-Hamiltonian matrix. As M itself is symmetric, the associated orthogonalization in (37) possesses a much simpler form, given by

$$\mathbf{q}_{j+1}\beta_j = M\mathbf{q}_j - \mathbf{q}_{j-1}\beta_{j-1} - \mathbf{q}_j\alpha_j - \sum_{i=1}^j J\mathbf{q}_i r_{ij}, \quad j = 1, \dots, k-1,$$
(39)

where

$$\alpha_j = \mathbf{q}_j^\top M \mathbf{q}_j, \quad \beta_j = \mathbf{q}_{j+1}^\top M \mathbf{q}_j,$$

and  $r_{ij}$  is given as in (38) which also equals zero in exact arithmetic. The equation (39) can be represented by the matrix form

$$MQ_k = Q_k T_k + J Q_k R_k + \mathbf{q}_{k+1} \beta_k \mathbf{e}_k^{\top}, \qquad (40)$$

where  $T_k = \text{tridiag}(\beta_{j-1}, \alpha_j, \beta_j) \in \mathbb{R}^{k \times k}$  is tridiagonal,  $R_k = [r_{ij}] \in \mathbb{R}^{k \times k}$  is upper triangular,  $Q_k = [\mathbf{q}_1 \cdots \mathbf{q}_k] \in \mathbb{R}^{2n \times k}$  is orthonormal and isotropic, and  $\mathbf{q}_{j+1}$  is a suitable vector satisfying

$$Q_k^{\top} \mathbf{q}_{j+1} = \mathbf{0} \quad \text{and} \quad (J Q_k)^{\top} \mathbf{q}_{j+1} = \mathbf{0}.$$
(41)

Deringer

According to (40) and (41), we get

$$Q_k^{\top} M Q_k = T_k \text{ and } \begin{bmatrix} Q_k^{\top} \\ (J Q_k)^{\top} \end{bmatrix} M \begin{bmatrix} Q_k & J Q_k \end{bmatrix} = \begin{bmatrix} T_k & \mathbf{0} \\ \mathbf{0} & T_k \end{bmatrix}.$$

Not only does the isotropic Lanczos process generate an orthonormal basis  $Q_k$  for the *k*-dimensional isotropic Krylov subspace, but also it splits the duplicate eigenvalues of M when M is projected onto the subspace generated by  $Q_k$ . So, to compute eigenvalues of  $T_k$ , and hence the Ritz values of M, we get each eigenvalue once, not twice. The critical essence of this numerical approach is that if we try to find, for instance, 2m distinct eigenvalues from M, then our method can produce 2m distinct eigenvalues from  $T_k$ , not a copy of m different eigenvalues.

Suppose that  $(\theta, \mathbf{z})$  is an eigenpair of  $T_k$ , i.e.,  $T_k \mathbf{z} = \theta \mathbf{z}$ . Then  $(\theta, Q_k \mathbf{z})$  is a Ritz pair of M. Then from (40) and (41) again, we have

$$\|(M - \theta I)Q_k \mathbf{z}\|_2 = \|(MQ_k - T_k Q_k)\mathbf{z}\|_2$$
  
=  $\|JQ_k R_k \mathbf{z} + \mathbf{q}_{k+1}\beta_k \mathbf{e}_k^\top \mathbf{z}\|_2 = \left\| \begin{bmatrix} R_k \mathbf{z} \\ \beta_k \mathbf{e}_k^\top \mathbf{z} \end{bmatrix} \right\|_2 \approx |\beta_k||\mathbf{e}_k^\top \mathbf{z}|.$  (42)

The approximately equal sign in (42) holds since  $R_k$  is an upper triangular matrix with tiny, or even zero, entries. Equation (42) provides us a simple and easy estimation of the residual  $||(M - \theta I)Q_k\mathbf{z}||_2$  as we obtain a Ritz pair  $(\theta, Q_k\mathbf{z})$  of M via the SHILA method.

*Remark 3* If we directly perform the classical Lanczos method to solve the CDSEP (32), the required iteration number may be more than the result of the SHILA method. This is because the double eigenvalues will hold each other up and tend to converge simultaneously. In contrast, the SHILA method can effectively exclude the inference of double eigenvalues and make efforts to the desired one. For more in-depth discussion, we refer to [32].

#### 5 Practical Implementations

The techniques of nonequivalence deflation and null-space free compression successfully reduce the original GEP (13) to a small-scale CDSEP (32). On the implementation, it is neither possible nor necessary to construct the inverse of deflating matrix  $\tilde{L}_C$  in (23a) and the compressed matrix  $\tilde{B}_1$  in (5.1).

In this section, we focus on how to efficiently compute the matrix-vector multiplication  $(\tilde{B}_1^{\top} \tilde{L}_C^{-1} \tilde{B}_1) \mathbf{q}$  for a given vector  $\mathbf{q}$  when we perform the SHILA iteration through the structures of "undeflated" matrices  $L_C$  and B themselves.

5.1 Explicit Eigendecomposition and Implicit Multiplication of  $\widetilde{B}$ 

As  $\widetilde{B}$  is a block-diagonal matrix consisting of  $\widetilde{D}$  in (23b), it is sufficient to determine the eigendecomposition of  $\widetilde{D}$ .

As in Sect. 4.1, we denote an ordered index set of the interior vertices by  $\mathcal{I}_{i}$  and use the notation  $\mathcal{I}_{b} = \{b_{1}, \ldots, b_{n_{b}}\}$  to indicate the corresponding ordered index set of boundary vertices. Let  $E_{0} \in \mathbb{R}^{n_{v} \times (n_{i}+1)}$  and  $E_{1} \in \mathbb{R}^{n_{v} \times (n_{b}-1)}$  be, respectively, defined by

$$E_0 = \begin{bmatrix} \mathbf{e}_{j_1} \cdots \mathbf{e}_{j_{n_{\perp}}} & \frac{1}{\sqrt{n_{\mathrm{b}}}} \mathbf{d} \end{bmatrix}, \quad j_1, \dots, j_{n_{\perp}} \in \mathcal{I}_{\perp},$$
(43a)

$$E_1 = \begin{bmatrix} \mathbf{b}_1 \cdots \mathbf{b}_{n_{\mathrm{b}}-1} \end{bmatrix},\tag{43b}$$

where **d** is the nonzero vector defined in (22) and columns of  $E_1$  are those  $n_v$ -vectors in (26). Then, by Lemma 2, we know that  $E_0$ ,  $E_1$  are column orthonormal matrices with  $E_0^{\top} E_1 = \mathbf{0}$ . Moreover, column vectors of the following enlarging matrices

$$\widetilde{B}_0 = \begin{bmatrix} E_0 & \mathbf{0} \\ \mathbf{0} & E_0 \end{bmatrix} \quad \text{and} \quad \widetilde{B}_1 = \begin{bmatrix} E_1 & \mathbf{0} \\ \mathbf{0} & E_1 \end{bmatrix}$$
(44)

are orthonormal eigenvectors of  $\tilde{B}$  corresponding to its eigenvalues 0 and 1, respectively (cf. Lemma 3). Based on the above deductions, we conclude that  $\tilde{B}$  is orthogonally diagonalizable,

$$\widetilde{B} = \begin{bmatrix} \widetilde{B}_0 & \widetilde{B}_1 \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{2(n_{\mathrm{b}}-1)} \end{bmatrix} \begin{bmatrix} \widetilde{B}_0^\top \\ \widetilde{B}_1^\top \end{bmatrix} = \widetilde{B}_1 \widetilde{B}_1^\top,$$

which is an explicit decomposition formula.

Although the low-rank compressed matrix  $\tilde{B}_1$  can be explicitly constructed, we still encounter another problem on the dense factorization. Nevertheless, we just need to know how to compute the product of  $\tilde{B}_1$  (or  $\tilde{B}_1^{\top}$ ) and a given vector when we use the SHILA method to solve the desired eigenpair. Fortunately, the specific structure of  $\tilde{B}_1$  provides an *implicitly* computational scheme of matrix-vector multiplication without explicitly generating  $\tilde{B}_1$  itself beforehand. From (44), it is thus sufficient to consider the matrix-vector multiplications:  $E_1\mathbf{q}$ and  $E_1^{\top}\mathbf{p}$  for given  $\mathbf{q} \in \mathbb{R}^{n_b-1}$  and  $\mathbf{p} \in \mathbb{R}^{n_v}$ . In order to make the expression clearer, we also adapt the reordering notations so that each column vector  $\mathbf{b}_k$  of  $E_1$  in (43b) is of the form (27). The actual computing can be achieved via appropriate index correspondences. For a fixed vector  $\mathbf{q} = [q_1 \cdots q_{n_b-1}]^{\top}$ ,  $E_1\mathbf{q}$  is an  $n_v$ -vector given by

where  $q_0 = 0$ . Similarly, the product of  $E_1^{\top}$  and a given vector  $\mathbf{p} = [p_1 \cdots p_{n_v}]^{\top}$  is the following  $(n_{\rm b} - 1)$ -vector

$$\left(E_{1}^{\top}\mathbf{p}\right)_{i} = \frac{1}{\sqrt{i(i+1)}} \sum_{k=1}^{l} \left(p_{k} - p_{i+1}\right), \quad i = 1, \dots, n_{\mathrm{b}} - 1.$$
(46)

# 5.2 The Inverse of $\tilde{L}_C$

When we apply the nonequivalence transformation (Sect. 4.1) and then reduce the deflating GEP (28) to the small-scale problem (32), at first glance, it seems definitely to modify  $L_C$  by a rank-2 update as presented in (23a). It is not advisable to construct the deflated matrix  $\tilde{L}_C$  as well as its inverse since the symmetric matrix  $(B\mathbb{I}_2)(B\mathbb{I}_2)^{\top}$  in (23), dependent on the boundary vertices, may be practically non-sparse. However, it is not necessary to actually compute  $\tilde{L}_C$  as we only need to be capable to find an equivalent way for solving the linear systems whose coefficient matrix is  $\tilde{L}_C$ .

To perform the SHILA method for solving the CDSEP (32) with the symmetric positive definite and skew-Hamiltonian matrix  $\tilde{B}_1^{\top} \tilde{L}_C^{-1} \tilde{B}_1$ , we must solve a linear system

$$\widetilde{L}_C \mathbf{p} = \mathbf{q} \tag{47}$$

with a given  $2n_{\nabla}$ -vector **q** in each step for the subspace expansion. Since  $\widetilde{L}_C$  is a rank-2 update of  $L_C$ , this motivates us to solve the linear system (47) based on the Sherman–Morrison–Woodbury formula (SMWF) [10]. Consider  $\widetilde{G} = G + UV^{\top}$  where  $G \in \mathbb{R}^{n \times n}$  is nonsingular,

 $U, V \in \mathbb{R}^{n \times r}$  are of rank  $r \ll n$  so that  $I_r + V^{\top} G^{-1} U$  is an  $r \times r$  invertible matrix. The SMWF suggests computing  $\tilde{G}^{-1}$  through the identity

$$\widetilde{G}^{-1} = (I_n - WV^{\top})G^{-1}$$
 with  $W = G^{-1}U(I_r + V^{\top}G^{-1}U)^{-1}$ . (48)

This formula is valid only if G is nonsingular while  $L_C$  fails to satisfy this requirement because of the zero row sum property [see (23a)]. However, we can rewrite  $\tilde{L}_C$  as

$$\widetilde{L}_{C} = L_{C}^{+} + \left[\frac{1}{n_{\rm b}}B\mathbb{I}_{2} - \mathbf{e}_{1} - \mathbf{e}_{n_{\rm v}+1}\right] \left[B\mathbb{I}_{2} \ \mathbf{e}_{1} \ \mathbf{e}_{n_{\rm v}+1}\right]^{\top},\tag{49}$$

where

$$L_C^+ \equiv L_C + \mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_{n_v+1} \mathbf{e}_{n_v+1}^\top = \begin{bmatrix} K^+ & S \\ -S & K^+ \end{bmatrix}, \quad K^+ = K + \mathbf{e}_1 \mathbf{e}_1^\top.$$
(50)

To contrast (49) with the SMWF formula (48), we see that  $n = 2n_{v}$ , r = 4,  $G = L_{C}^{+}$ ,  $V = \begin{bmatrix} B\mathbb{I}_{2} \ \mathbf{e}_{1} \ \mathbf{e}_{n_{v}+1} \end{bmatrix}$  and W can be expressly formulated as follows. Let  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n_{v}}$  satisfy  $L_{C}^{+}\begin{bmatrix}\mathbf{x}\\\mathbf{y}\end{bmatrix} = \frac{1}{n_{b}}\begin{bmatrix}\mathbf{d}\\\mathbf{0}\end{bmatrix}$ , where  $\mathbf{d}$  is defined as in (22). Set  $\chi = \frac{1}{n_{b}}(\mathbf{d}^{\top}\mathbf{x}+1)$ ,  $\eta = \frac{1}{n_{b}}(\mathbf{d}^{\top}\mathbf{y})$ ,  $\sigma = (x_{1}^{2} + y_{1}^{2})^{-1}n_{b}x_{1}$ ,  $\tau = (x_{1}^{2} + y_{1}^{2})^{-1}n_{b}y_{1}$  with  $x_{1} = \mathbf{e}_{1}^{\top}\mathbf{x}$  and  $y_{1} = \mathbf{e}_{1}^{\top}\mathbf{y}$ . Then

$$W = \frac{1}{n_{\rm b}} \begin{bmatrix} \mathbf{1}_{n_{\rm v}} & \mathbf{0} & \mathbf{h} & \mathbf{k} \\ \mathbf{0} & \mathbf{1}_{n_{\rm v}} & -\mathbf{k} & \mathbf{h} \end{bmatrix},\tag{51}$$

where  $\mathbf{h} = \sigma \mathbf{x} + \tau \mathbf{y} - (\sigma \chi + \tau \eta) \mathbf{1}_{n_{v}}$  and  $\mathbf{k} = \tau \mathbf{x} - \sigma \mathbf{y} - (\tau \chi - \sigma \eta) \mathbf{1}_{n_{v}}$ .

This representation views  $\tilde{L}_C$  as a rank-4 update of  $L_C^+$ .  $L_C^+$  itself is a simple perturbation, virtually no-cost, on the first diagonal element of  $L_C$  as well as its dual. From (50), one can see that  $L_C^+$  completely preserves the structure and sparsity of  $L_C$ , and, in general,  $L_C^+$  will be a nonsingular matrix. Therefore, we present a feasible alternative for applying the SMWF to  $\tilde{L}_C$  for solving (47).

*Remark* 4 The low-rank update formula (50) can be written in a more general form

$$K^+ = K + \mathbf{e}_k \mathbf{e}_k^{\mathsf{T}}, \quad 1 \le k \le n_{\mathsf{v}}.$$

In other words, we can select suitable vertex index k to destroy the zero row sum property of K so that  $K^+$  (and hence  $L_C^+$ ) becomes a nonsingular matrix. Fixed  $1 \le k \le n_{v}$ , the  $2n_{v}$ -vector  $\mathbf{e}_1$  and its dual one  $\mathbf{e}_{n_v+1}$  in Eqs. (49)–(51), of course, will be changed accordingly. The resulting matrix  $L_C^+$  and the original one  $L_C$  always have the same structure and sparsity as shown in (49) and (50). So one can adaptively choose the perturbation term to construct an invertible matrix  $L_C^+$  so that SMWF formula (48) is applicable for solving the linear system (47).

# 5.3 SHILA for Spectral Conformal Parameterizations

Algorithm 5.1 summarizes the SHILA method for solving the smallest positive eigenvalue and corresponding eigenvector of the GEP (13) based on a null-space free compression technique to the deflated matrix pair ( $\tilde{L}_C$ ,  $\tilde{B}$ ) in (23). We then explain some key points for our MATLAB implementation of Algorithm 5.1.

(i) All matrix-vector multiplications, including  $\tilde{B}_1$  in (31) and its transpose  $\tilde{B}_1^{\top}$ , can be performed by multiply-add operations on vectors with appropriate indices [cf. (44)–(46)]. So, it does not require any extra cost to construct and store these matrices.

#### Algorithm 5.1 SHILA Procedure for Spectral Conformal Parameterizations

**Input:** The matrix  $L_C$  in (7); a random unit vector  $\mathbf{q}_1$ ; the maximum iteration *maxit* and the tolerance *tol*. **Output:**  $(\lambda, \mathbf{f})$  where  $\lambda$  is the smallest positive eigenvalue of the GEP (13) and  $\mathbf{f}$  is the associated eigenvector.

1: % Initialization  
2: Set 
$$Q_1 = q_1$$
 and  $P_0 = []$ ;  
3:  $L_C \leftarrow L_C^+$  by (50);  
4: % The SHILA procedure  
5: for  $j = 1, 2, ..., maxi do$   
6: % Compute  $q = \tilde{B}_1^- \tilde{L}_C^{-1} \tilde{B}_1 q_j$   
7: Compute  $q = \tilde{B}_1^- \tilde{L}_C^{-1} \tilde{B}_1 q_j$ ;  
8: Solve  $\tilde{L}_C \mathbf{p}_j = t$  by SMWF with (49)–(50);  
9: Set  $P_j = [P_{j-1}, \mathbf{p}_j] = \tilde{L}_C^{-1} \tilde{B}_1 Q_j$ ;  
10: Compute  $q = \tilde{B}_1^- \mathbf{p}_j = \begin{bmatrix} \tilde{E}_1^- 0 \\ 0 & E_1^- \end{bmatrix} \mathbf{p}_j$  implicitly by (46);  
11: % Orthogonalization process  
12: Compute  $\alpha_j = \mathbf{q}_j^- \mathbf{q}_i$ ;  
14: if  $j > 1$  then  
15:  $\mathbf{q} \leftarrow \mathbf{q} - \alpha_j q_j$ ;  
16: end if  
17: % Isotropic orthogonalization process  
18: Compute  $r = (JQ_j)^+ \mathbf{q}$ , where  $J$  is the matrix (8);  
19:  $\mathbf{q} \leftarrow \mathbf{q} - (JQ_j)r$ ;  
20: % Full reorthogonalization process  
21: Compute  $r = Q_j^- \mathbf{q}_i$ ;  
22:  $\mathbf{q} \leftarrow \mathbf{q} - Q_j \mathbf{r}_i$ ;  
33: % Compute the  $j + 1$  Lanczos vector  $\mathbf{q}_{j+1}$   
24: Compute the  $j + \mathbf{q}/\beta_j$  and  $Q_{j+1} = [Q_j, \mathbf{q}_{j+1}]$ ;  
25: Set  $\mathbf{q}_{j+1} = \mathbf{q}/\beta_j$  and  $T_j := \begin{bmatrix} \frac{T_{j-1}}{0} & \frac{\mathbf{0}}{\beta_{j-1}} \\ 0 & \beta_{j-1} & \alpha_j \end{bmatrix}$ ,  $j \ge 3$ ;  
28: % Check residual  
29: if  $|\beta_j||\mathbf{e}_j^- \mathbf{z}| < tol then$   
30: Set  $a \in a^{-1}$  and  $\mathbf{f} = \lambda P_j z$ ;  
return  $(\lambda, f)$ 

(ii) In order to improve the efficiency of solving the linear system in line 8 of Algorithm 5.1, we first compute the Cholesky factorization of  $L_C^+$  by the MATLAB function chol, before for-loop procedure of Algorithm 5.1, and use the resulting factorization together with SMWF as well as the MATLAB built-in function *matrix left division* (i.e., \) to solve the linear system in line 8 at each isotropic Lanczos step. If the matrix size is extremely large, any suitable iterative method can be a feasible alternative.

- (iii) It is essential to prevent the loss of isotropicity and orthogonality of Lanczos vectors. To this end, we perform the isotropic reorthogonalization and the full reorthogonalization as shown in lines 17–19 and 20–22 of Algorithm 5.1, respectively.
- (iv) To seek the eigenpair of the small-scale  $T_j$  with largest magnitude (line 27), we call the eig function to accomplish this task.
- (v) To determine the stopping criterion *tol* in line 29 of Algorithm 5.1, we propose a lower bound estimation of the residual  $||L_C \mathbf{f} \lambda B \mathbf{f}||_2$  as follows.

**Lemma 4** Let  $(\lambda^{-1}, \mathbf{s}_1)$  be an approximate eigenpair of the CDSEP (32) which is the candidate for approximating the largest real positive eigenvalue and its associated eigenvector. Then

$$(\lambda \| \widetilde{L}_C^{-1} \|_2^{-1}) \left\| \widetilde{B}_1^\top \widetilde{L}_C^{-1} \widetilde{B}_1 \mathbf{s}_1 - \frac{1}{\lambda} \mathbf{s}_1 \right\|_2 \le \| L_C \mathbf{f} - \lambda B \mathbf{f} \|_2,$$
(52)

where  $\mathbf{f}$  is defined as in (35).

*Proof* By (33)–(35) and (29), we get

$$\begin{split} \left\| \widetilde{B}_{1}^{\top} \widetilde{L}_{C}^{-1} \widetilde{B}_{1} \mathbf{s}_{1} - \frac{1}{\lambda} \mathbf{s}_{1} \right\|_{2} &= \left\| \widetilde{B}_{1}^{\top} (\widetilde{L}_{C}^{-1} \widetilde{B} \mathbf{f} - \frac{1}{\lambda} \mathbf{f}) \right\|_{2} \\ &\leq \frac{1}{\lambda} \| \widetilde{L}_{C}^{-1} \|_{2} \| \widetilde{L}_{C} - \lambda \widetilde{B} \mathbf{f} \|_{2} = \frac{1}{\lambda} \| \widetilde{L}_{C}^{-1} \|_{2} \| L_{C} - \lambda B \mathbf{f} \|_{2}. \end{split}$$

Note that  $\|\widetilde{L}_{C}^{-1}\|_{2}^{-1}$  equals to the smallest positive eigenvalue of  $L_{C}$ , the right hand side of (52) can be approximated by  $\lambda^{2} \|\widetilde{B}_{1}^{\top} \widetilde{L}_{C}^{-1} \widetilde{B}_{1} \mathbf{s}_{1} - \frac{1}{\lambda} \mathbf{s}_{1} \|_{2}$ . Since the smallest positive eigenvalue  $\lambda$  of the GEP (13) can be close to zero, we expect that the lower bound of  $\|L_{C}\mathbf{f} - \lambda B\mathbf{f}\|_{2}$  in (52) can be very small when  $\lambda$  is very close to the desire target. Lemma 4 therefore provides us an estimation to determine the stopping criterion *tol* in line 29 Algorithm 5.1 [(see also (42)]. In each SHILA step, we can adaptively set *tol* so that

$$\lambda^2 \|\widetilde{B}_1^\top \widetilde{L}_C^{-1} \widetilde{B}_1 \mathbf{s}_1 - \frac{1}{\lambda} \mathbf{s}_1\|_2 \le \lambda^2 tol \approx eps.$$

In our implementations, we observe that almost all (approximate) eigenvalues are  $< 10^{-5}$  (see Table 1) and hence the lower bound on the left hand side of (52) is approximate to  $\|\tilde{B}_1^{\top}\tilde{L}_C^{-1}\tilde{B}_1\mathbf{s}_1 - \frac{1}{\lambda}\mathbf{s}_1\|_2$  times a factor  $< 10^{-10}$ . Therefore, for simplicity, we take  $tol = 10^{-5}$  to be a fixed tolerance as the stopping criterion of SHILA algorithm and expect that the accuracy for the residual of the returning Ritz pair ( $\lambda$ ,  $\mathbf{f}$ ) in line 30 of Algorithm 5.1 will be sufficiently precise.

(vi) Finally, and most importantly, we highlight the statements in lines 9 and 30 of Algorithm 5.1 which need further explanation. Suppose we obtain an desired eigenpair ( $\theta$ , z) of  $T_k$  at step k, i.e., ( $\theta$ , z) satisfies the stopping criterion (42) in line 29 of Algorithm 5.1. Then, ( $\theta$ ,  $Q_k z$ ) is a Ritz pair of the CDSEP (32). With the aid of (35), we know that

$$\lambda = \theta^{-1} \quad \text{and} \quad \mathbf{f} = \lambda L_C^{-1} B_1 Q_k \mathbf{z} \tag{53}$$

are the approximate eigenvalue and the associated eigenvector of the original eigenproblem (13). Therefore, to get the desired eigenvector, (53) indicates that we need to perform two matrix-vector multiplications and solve one linear system. But this is only a theoretical expression. In practice, the pre-stored matrix  $P_k$  in line 9 of Algorithm 5.1 collects the computational results of  $\tilde{L}_C^{-1}\tilde{B}_1Q_k$  in our past effort in lines 7 and 8. So, to transform **z** back to the vector **f** in (53), we actually just need to compute one matrixvector multiplication together with a scalar product as shown in line 30 of Algorithm 5.1, without extra cost for solving a linear system.

# 6 Numerical Experiments

We demonstrate the efficiency and accuracy of our novel numerical technique, SHILA (Algorithm 5.1), to solve the new derived CDSEP (32) for the computation of DCPs of various mesh models.

# 6.1 Performance

All numerical demonstrations in this work were carried out using MATLAB R2013a on a MacBook Pro Retina with 2.6 GHz Intel Core i5 processor and 8 GB of RAM. The maximum iteration *maxit* and the stopping tolerance *tol* in Algorithm 5.1 are taken as *maxit* = 30 and  $tol = 10^{-5}$  [see (v) of Sect. 5.3], respectively.

In order to show the advantages of SHILA, we additionally apply the *classical Lanczos method* to solve our CDSEP (32). Almost all of the procedures of its implementations are the same as those of the SHILA method (as described in Sect. 5.3) except for the isotropic orthogonalization process in lines 17–19 of Algorithm 5.1. Moreover, we also consider the performance of the modified GEP (17) in [31] and the Schur complement reduction (19) in [1]. To solve the modified GEP (17), we call the eigs function from the MATLAB library to find the largest eigenvalue and associated eigenvector of (17) with a function handle to compute the matrix-vector product on the left-hand side of (17) without explicitly forming this matrix. For the approach introduced in [1], we first compute the coefficient matrix of the Schur complement as in (19) (via mldivide) and subsequently solve the reduced eigenvalue problem by calling eigs. Note that the reduced system (19) still has double eigenvalues including the zeros, so we call eigs to compute 4 eigenvalues with the smallest magnitude—two of them are 0 and the others are (theoretically) identical to the smallest positive eigenvalue of the GEP (13).

To express numerical results of these four methods, we denote  $SL_{CDSEP}$ ,  $L_{CDSEP}$ ,  $E_{MGEP}$  and  $SC_{RGEP}$  as follows:

- SL<sub>CDSEP</sub>: Solving the CDSEP (32) by the SHILA method (Algorithm 5.1).
- $L_{CDSEP}$ : Solving the CDSEP (32) by the classical Lanczos method.
- $E_{MGEP}$ : Solving the modified GEP (17) by the MATLAB function eigs.
- $SC_{RGEP}$ : Solving the reduced GEP (19) by the Schur complement method.

To quantitatively measure the conformality, we adapt the *quasi-conformal* (QC) distortion [35]. The QC distortion factor is computed per mesh triangle face as the ratio  $\frac{\Gamma}{\gamma}$ , where  $\Gamma$  and  $\gamma$  are larger and smaller eigenvalues of the Jacobian of the map. The ideal conformality is 1, larger values indicate worse conformality.

Table 1 shows the computational time of these methods. For  $SL_{CDSEP}$  and  $L_{CDSEP}$ , we additionally present the individual iteration numbers. Moreover, as we obtain the desired eigenpair  $(\lambda, \mathbf{f})$  from  $SL_{CDSEP}$ ,  $L_{CDSEP}$ ,  $E_{MGEP}$  or  $SC_{RGEP}$ , we compute the residual  $||L_C \mathbf{f} - \lambda B\mathbf{f}||_2$  and  $||\mathbf{f}^\top B \mathbb{I}_2||_2$  for each approach as shown in Fig. 1. Note that if  $\mathbf{f}$  is the eigenvector computed from  $E_{MGEP}$ , then, in general,  $||\mathbf{f}^\top B \mathbf{f} - 1||_2 \gg 0$ . So, in this case,  $\mathbf{f}$  should be further normalized by  $\frac{\mathbf{f}}{\sqrt{\mathbf{f}^\top R \mathbf{f}}}$ .

For the sake of brevity, we only present 10 representative figures (Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10, 11) to show the parameterization results among these models (all parameterizations can be found on the website of the first author <sup>3</sup>). In each figure, (a) is the original triangular mesh model, (b) shows the embedded parameter domain of SCP with color coded QC distortion, and (c) presents the SCP result with texture mapping.

<sup>&</sup>lt;sup>3</sup> http://jupiter.math.nctu.edu.tw/~wqhuang/shila/shila\_scp.pdf



Fig. 1 Residuals and *B*-orthogonality for  $SL_{CDSEP}$ ,  $L_{CDSEP}$ ,  $E_{MGEP}$  and  $SC_{RGEP}$ . a Mesh Models versus Residuals, b Mesh Models versus *B*-orthogonality



# 6.2 Discussions

• *Efficiency* From Table 1, we observe that the  $SC_{RGEP}$  scheme is competitive with other methods only when the number of boundary vertex  $n_b$  is moderate. As  $n_v$  and  $n_b$  increase, especially for large  $n_b$ ,  $SC_{RGEP}$  approach takes more and more time as well as memory to solve the matrix equation (21) and to store the computed results. Even in the example



Fig. 4 The Julius Caesar model



Fig. 6 The Foot model

on the Uzzano model, the  $SC_{RGEP}$  method was unable to solve the matrix equation (21) in 30 min and it eventually ran out of memory after 2,331 s. In contrast,  $E_{MGEP}$  has much better efficiency than the  $SC_{RGEP}$  approach. But, with the increasing numbers of  $n_i$  and  $n_b$ , we see that, once again,  $E_{MGEP}$  spends even two to four times more CPU time than those of these two Lanczos-type methods. The isotropic orthogonalization process is the significant difference between SHILA and the classical Lanczos method. First of all, the CPU time costs in Table 1 show that the isotropic orthogonalization process (in lines 17–19 of Algorithm 5.1) does not have a significant amount of time spent even



Fig. 8 The Hand model

when the iteration number of  $SL_{CDSEP}$  is equal to that of  $L_{CDSEP}$ . Moreover, according to the numbers of iterations, one can see that SHILA can absolutely remove the influence of duplicate eigenvalues, while the iteration numbers of  $L_{CDSEP}$  show that the classical Lanczos method can suffer the impact of double eigenvalues. Our experiments reveal that the iteration numbers of  $SL_{CDSEP}$  are less than five to nine iterations of  $L_{CDSEP}$  procedure. This indicates that SHILA can efficiently improve the convergence rate and reduce the required CPU time.

• Accuracy From Fig. 1, we observe that  $\mathbb{E}_{MGEP}$  lose the precision of the residuals  $||L_C \mathbf{f} - \lambda \mathbf{f}||_2$ , and the *B*-orthogonality ( $\mathbf{f}^\top B \mathbb{I}_2 = 0$ ). Since (17) is a singular GEP and almost any perturbation of a singular pencil will turn to a regular one (nearly singular), the eigs function can be used to compute the desired eigenpair of the nearly singular pencil corresponding to (17). Nevertheless, to compute the eigenpairs of a nearly singular pencil is very sensitive [3,30]. As shown in Fig. 1b, we see that the eigenvector  $\mathbf{f}$  of (17), computed by eigs, cannot possess the *B*-orthogonality. Consequently,  $\mathbb{E}_{MGEP}$  has only the accuracy of residuals between  $10^{-6} \sim 10^{-9}$  (see Fig. 1a). On the other hand,  $SL_{CDSEP}$  and  $L_{CDSEP}$  always have the same accuracy and are almost more accurate than  $SC_{RGEP}$ 



1.1 1.15

1.2 1.25

(b)

1.3

(c)





Fig. 11 The Vase Lion model

Model	n <sub>v</sub>	nb	λ	Time (#Its) [seconds]			
				SL <sub>CDSEP</sub>	L <sub>CDSEP</sub>	E <sub>MGEP</sub>	SC <sub>RGEP</sub>
Susan	5,161	321	4.8045e-6	0.1 (7)	0.1 (7)	0.1	0.7
Fandisk	6,699	450	2.6814e-6	0.1 (8)	0.1 (8)	0.1	0.8
Saddle	8,321	256	3.3305e-7	0.1 (6)	0.1 (11)	0.4	1.0
Foot	10,211	454	1.9270e-6	0.1 (8)	0.1 (13)	0.3	1.1
Gargoyle	10,229	456	1.6899e-5	0.1 (10)	0.1 (10)	0.4	1.2
Beetle	15,053	375	5.2974e-7	0.2 (7)	0.3 (12)	0.3	6.5
Sophie	15,282	562	2.1015e-6	0.2 (7)	0.3 (13)	0.3	4.2
Pipes	20,304	64	1.0470e-4	0.3 (6)	0.3 (6)	1.1	3.7
Dino	24,605	1,248	1.0443e-6	0.3 (10)	0.4 (17)	0.9	7.8
Bunny	35,190	927	1.9835e-6	0.6 (10)	0.6 (10)	1.9	15.0
Hand	37,234	1,508	2.0250e-7	0.6 (9)	0.8 (15)	1.7	26.0
Camel	402,40	2,334	5.4112e-7	0.7 (11)	1.0 (20)	2.1	45.4
Vase-Lion	178,491	625	6.8350e-6	2.6 (9)	2.6 (9)	3.8	59.2
Isis	188,144	1,977	6.4689e-8	4.4 (10)	5.0 (15)	11.6	679
Planck	199,169	293	2.7225e-6	4.3 (7)	4.3 (7)	11.8	34.8
Chinese-Lion	255,284	928	7.5866e-6	5.7 (11)	5.7 (11)	9.2	371
Caesar	387,900	1,634	9.2161e-7	11.3 (9)	13.4 (16)	48.6	1,227
Bimba	423,713	1,308	8.2092e-8	7.6 (8)	8.8 (13)	23.8	1,180
Uzzano	946,451	2,581	1.4856e-7	36.3 (10)	61.8 (17)	65.9	_

Table 1 CPU time and number of iterations. Note that for the Uzzano model,  $SC_{RGEP}$  did not complete and encountered the "out of memory" error in 2,331 s after

(about 1 significant digit). Moreover, both  $SL_{CDSEP}$  and  $L_{CDSEP}$  have high precision on the requirements of the *B*-orthogonality.

In summary,  $SL_{CDSEP}$  has much better efficiency than  $SC_{RGEP}$  and better than  $L_{CDSEP}$  and  $E_{MGEP}$ . For accuracy,  $SL_{CDSEP}$  and  $L_{CDSEP}$  present high numerical accuracy. The accuracy of  $SL_{CDSEP}$  is much better than  $E_{MGEP}$  and is better than  $SC_{RGEP}$ . These data advocate the efficiency and accuracy of  $SL_{CDSEP}$ , *a new derived* CDSEP (32) *together with the novel* SHILA *algorithm*, as a fundamental tool for computing SCPs.

# 7 Conclusions

Spectral methods are not new in computer graphics and geometry processing, and have been developed to solve a diversity of problems. For an up-to-date survey on this topic, we refer to [43]. Since the potential quantities, such as eigenvalues and eigenvectors, of a spectral method are the primary factors in solving such problems, how to accurately and efficiently dig out these pieces of eigeninformation is always an important issue.

Spectral conformal parameterization (SCP) [31] is one of the various applications in spectral mesh processing. To compute a conformal mesh parameterization, SCP suggests to compute the eigenvector the GEP (13) corresponding to its smallest positive eigenvalue. By inspecting the particular matrix structures of the pair ( $L_C$ , B) in (13), we can apply appro-

priate nonequivalence deflation and null-space free compression techniques to transform this eigenproblem into the small-scale CDSEP (32) with a symmetric positive definite skew-Hamiltonian operator. We derive an explicit representation for the coefficient matrices of the CDSEP (32) and propose an implicit technique for practical implementation. Furthermore, We develop a novel SHILA algorithm for solving the CDSEP (32).

Numerical experiments show that, compared with the numerical implementations in [31] and [1], the new derived CDSEP together with the SHILA algorithm can accurately and robustly compute the parameterization results of SCP.

Acknowledgments Some of the mesh models are available on the websites CCGL (Susan and Sophie); Unwrapped meshes (Bunny, Camel, Dino, Fandisk, Hand, Foot and Isis), and Project page of ARAP (Beetle and Gorgyle). The remaining mesh models are courtesy of AIM@SHAPE Shape Repository. The first and third authors would like to acknowledge the support from the National Science Council and the National Centre for Theoretical Sciences in Taiwan. They also like to thank the ST Yau Centre at the National Chiao Tung University for the support.

## References

- Alexa, M., Wardetzky, M.: Discrete Laplacians on general polygonal meshes. In: ACM SIGGRAPH 2011 Papers, SIGGRAPH '11, pp. 102:1–102:10. ACM, New York, NY, USA (2011)
- Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., Sorensen, D.: LAPACK Users' Guide, 3rd edn. Society for Industrial and Applied Mathematics, Philadelphia, PA (1999)
- Bai, Z., Demmel, J., Dongarra, J., Ruhe, A., van der Vorst, H.: Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide. Society for Industrial and Applied Mathematics, Philadelphia, PA (2000)
- Balay, S., Adams, M.F., Brown, J., Brune, P., Buschelman, K., Eijkhout, V., Gropp, W.D., Kaushik, D., Knepley, M.G., McInnes, L.C., Rupp, K., Smith, B.F., Zhang, H.: PETSc Web page. (2014). http://www. mcs.anl.gov/petsc
- Cohen-Steiner, D., Desbrun, M.: Hindsight: LSCM and DCP are one and the same. (2002). http://www.geometry.caltech.edu/pubs/CD02.pdf
- Desbrun, M., Meyer, M., Alliez, P.: Intrinsic parameterizations of surface meshes. Comput. Graph. Forum 21(3), 209–218 (2002)
- 7. Fiedler, M.: Algebraic connectivity of graphs. Czechoslov. Math. J. 23(2), 298–305 (1973)
- Floater, M., Hormann, K.: Surface parameterization: a tutorial and survey. In: Dodgson, N., Floater, M., Sabin, M. (eds.) Advances in Multiresolution for Geometric Modelling, Mathematics and Visualization, pp. 157–186. Springer, Berlin (2005)
- Floater, M.S.: Parametrization and smooth approximation of surface triangulations. Comput. Aided Geom. Des. 14(3), 231–250 (1997)
- Golub, G.H., Van Loan, C.F.: Matrix Computations, 4th edn. The Johns Hopkins University Press, Baltimore, MD (2012)
- Gortler, S.J., Gotsman, C., Thurston, D.: Discrete one-forms on meshes and applications to 3D mesh parameterization. Comput. Aided Geom. Des. 23(2), 83–112 (2006)
- Gotsman, C., Gu, X., Sheffer, A.: Fundamentals of spherical parameterization for 3D meshes. ACM Trans. Graph. 22(3), 358–363 (2003)
- Gu, D., Luo, F., Yau, S.T.: Fundamentals of computational conformal geometry. Math. Comput. Sci. 4(4), 389–429 (2010)
- Gu, X., Yau, S.T.: Global conformal surface parameterization. In: Proceedings of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing, SGP '03, pp. 127–137. Eurographics Association, Aire-la-Ville, Switzerland (2003)
- Gu, X., Yau, S.T.: Computational Conformal Geometry. International Press, Somerville, MA, Higher Education Press, Beijing (2008)
- Gu, X., Zeng, W., Luo, F., Yau, S.T.: Numerical computation of surface conformal mappings. Comput. Methods Funct. Theory 11(2), 747–787 (2012)
- Hernandez, V., Roman, J.E., Vidal, V.: SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems. ACM Trans. Math. Softw. 31(3), 351–362 (2005)

- Hoppe, H., Praun, E.: Shape compression using spherical geometry images. In: Dodgson, N., Floater, M., Sabin, M. (eds.) Advances in Multiresolution for Geometric Modelling, Mathematics and Visualization, pp. 27–46. Springer, Berlin (2005)
- Hormann, K.: Theory and Applications of Parameterizing Triangulations. Ph.D. thesis. University of Erlangen, Erlangen, Germany (2001)
- Hormann, K., Polthier, K., Sheffer, A.: Mesh Parameterization: Theory and practice. SIGGRAPH Asia courses (2008)
- Jin, M., Kim, J., Luo, F., Gu, X.: Discrete surface Ricci flow. IEEE Trans. Vis. Comput. Graph. 14(5), 1030–1043 (2008)
- Jin, M., Wang, Y., Yau, S.T., Gu, X.: Optimal global conformal surface parameterization. In: IEEE Visualization, pp. 267–274 (2004)
- Kharevych, L., Springborn, B., Schröder, P.: Discrete conformal mappings via circle patterns. ACM Trans. Graph. 25(2), 412–438 (2006)
- Lehoucq, R.B., Sorensen, D.C., Yang, C.: ARPACK Users' Guide: Solution of Large Eigenvalue Problems with Implicitly Restarted Arnoldi Methods. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1998)
- Lévy, B., Petitjean, S., Ray, N., Maillot, J.: Least squares conformal maps for automatic texture atlas generation. ACM Trans. Graph. 21(3), 362–371 (2002)
- 26. Lin, W.W.: Beiträge zur numerischen behandlung des allgemeinen eigenwertproblems  $Ax = \lambda Bx$ . Ph.D. thesis. Universität Bielefeld, Bielefeld, Germany (1985)
- Lin, W.W.: On reducing infinite eigenvalues of regular pencils by a nonequivalence transformation. Linear Algebra Appl. 78, 207–231 (1986)
- Marchandise, E., Remacle, J.F., Geuzaine, C.: Quality surface meshing using discrete parametrizations. In: Quadros, W. (ed.) Proceedings of the 20th International Meshing Roundtable, pp. 21–39. Springer, Berlin (2012)
- Mehrmann, V., Watkins, D.: Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils. SIAM J. Sci. Comput. 22(6), 1905–1925 (2001)
- Moler, C., Stewart, G.: An algorithm for generalized matrix eigenvalue problems. SIAM J. Numer. Anal. 10(2), 241–256 (1973)
- Mullen, P., Tong, Y., Alliez, P., Desbrun, M.: Spectral conformal parameterization. Proceedings of the Symposium on Geometry Processing, SGP '08, pp. 1487–1494. Eurographics Association, Aire-la-Ville, Switzerland (2008)
- Parlett, B.N.: The Symmetric Eigenvalue Problem. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1998)
- Pinkall, U., Polthier, K.: Computing discrete minimal surfaces and their conjugates. Exp. Math. 2(1), 15–36 (1993)
- Ray, N., Li, W.C., Lévy, B., Sheffer, A., Alliez, P.: Periodic global parameterization. ACM Trans. Graph. 25(4), 1460–1485 (2006)
- Sander, P.V., Snyder, J., Gortler, S.J., Hoppe, H.: Texture mapping progressive meshes. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '01, pp. 409–416. ACM, New York, NY, USA (2001)
- Sheffer, A., Lévy, B., Mogilnitsky, M., Bogomyakov, A.: ABF++: fast and robust angle based flattening. ACM Trans. Graph. 24(2), 311–330 (2005)
- Sheffer, A., Praun, E., Rose, K.: Mesh parameterization methods and their applications. Found. Trends Comput. Graph. Vis. 2(2), 105–171 (2006)
- Sheffer, A., de Sturler, E.: Parameterization of faceted surfaces for meshing using angle-based flattening. Eng. Comput. 17(3), 326–337 (2001)
- Tisseur, F.: Stability of structured Hamiltonian eigensolvers. SIAM J. Matrix Anal. Appl. 23(1), 103–125 (2001)
- Van Loan, C.F.: A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix. Linear Algebra. Appl. 61, 233–251 (1984)
- Yang, Y.L., Guo, R., Luo, F., Hu, S.M., Gu, X.: Generalized discrete Ricci flow. Comput. Graph. Forum 28(7), 2005–2014 (2009)
- Zayer, R., Lévy, B., Seidel, H.P.: Linear angle based parameterization. In: Proceedings of the Fifth Eurographics Symposium on Geometry Processing, SGP '07, pp. 135–141. Eurographics Association, Aire-la-Ville, Switzerland (2007)
- Zhang, H., Van Kaick, O., Dyer, R.: Spectral mesh processing. Comput. Graph. Forum 29(6), 1865–1894 (2010)